

Weekly Report

Done

1. 与彭老师巫老师开会以及一些后续查资料以及讨论. 这个题可以作如下概括: 一个同时考虑序列间关系以及序列自身演变 **patterns** 的事件序列可视化方法, 主要贡献将是一个主视图(或者两个可以平滑的在两种模式(序列间和序列内)切换的可视化方法). 这个问题虽可以算一个新问题, 但是仍比较抽象, 放到以后去想.
2. 与陈老师的讨论及查阅相关资料: 主旨是从数据直接生成视频. 借鉴 2009 年的基于 **event graph** 的 **digital storytelling** 思想和手法. 下面列出一些讨论或者思考的结果:
 - a. 保持 **event graph** 不变.
 - b. 对于 **viewpoint** 等参数的重定义. 剪辑电影时, 同样有镜头内转换与镜头间的衔接问题. 一方面, 从信息可视化角度的重新类比, 比如 **zoom in/out**, 如从散点到柱形图, 数据方阵的转置, 改变 **encoding**, 改变 **layout**, 颜色的过渡(以上种种可以请教梅博), 另一方面则可借鉴电影相关技术, 如转场\组接\蒙太奇手法等, 这些也同样可作为一些指导 **event graph** 生成路径的准则
 - c. **event graph** 生成路径的准则可以用一个优化目标函数去定义, 让算法自动搜索出较好的路径. 这个可以联想今年的 **infovis best paper** 用的 **ASP solver**, 当然也可以用其他图上面的算法进行操作. 这里的优化目标函数, 可结合数据探索的指标(当前数据覆盖率, 节点转移导致的认知成本, 数据类型, 数据之间的联系(时间的空间的))等等.
 - d. **VIS17 Poster, DataTours: A Data Narratives Framework** 一文, 他们把 **graph** 做的很复杂精妙, 但正如梅博所说, 我们毕竟还是要"剪辑"视频的, 诸如 **annotation** 这样的操作可以放在"剪"的阶段作为微调. 这里是要权衡一下 **graph** 设计与这种后期剪辑的; 后期剪辑我们当然也可以尽可能的智能, 比如可以推荐一些 **annotation**.
3. 一些电影相关知识的了解, 做了一些简单类比. 虽然比较细但也很实际的美学意义, 同时类比到数据可视化领域也有一定的道理所在(比如时间分配, 远近切换).
4. 整理了一些网易游戏数据; 以及本科同学期末汇报和报告的一些梳理和指导

To Do

1. 找找电影相关老师再做一些简单了解(我女朋友的本科毕设导师), 看看有没有一些启发性的概念可以借用. 从我粗略了解看, 蒙太奇这个词是电影动画相关常见技巧, 他们本身也是对信息在时间空间含义上整合的过程, 我希望再有一些了解.
2. 和老师同学们再细化一下上面想法.

论文阅读

1. 泛读两篇 word embedding 可解释性的工作. SPINE: SParse Interpretable Neural Embeddings, 理解 NLP 组对于 interpretable 的定义以及测试方法.
2. Storytelling via Navigation: A novel approach to annotation for sci-vis: 针对用户是 domain expert; 用户像往常一样探索一个体数据集; 探索过程中, 用户发现重要的 views, 并给每个 view 写一点 description. 根据这些来判断哪些重要帧, 同时用户可以自定义重要的帧, 最终合成视频用户探索时, 一般一次只改变一个参数; 在 nail down 到细节之前, 必有 trial-and-err 过程
3. Modeling Individual Cyclic Variation in Human Behavior 用隐马尔科夫来发现周期性模式, 人之间差异和个人自己的 variation 的方法. 代码在 github 有, 如果对于那种有很强周期性的数据会很适用.
4. BigIN4: Instant, Interactive Insight Identification for Multi-Dimensional Big Data: 用 bayesian 网络加速查询结果(估算). 在查询时, 为了快速得到 insight 往往需要做一些 data cube 加速. 有时候正好这个 cube 没被预计算那就要花很多时间实时计算.

工作时间

平时 60, 周末一共 9 小时. 总共 69 小时.